DOCUMENT RESUME

ED 161 942 TM 007 977

AUTHOR Pillos, Rita M.; Magoon, A. Jon
TITLE Evaluation Acceptance in Elementary School Teachers: A Construct Validation and Description of the Meaning of Standardized Achievement Testing.
PUB DATE [Mar 78]
NOTE 38p.; Paper presented at the Annual Meeting of the American Educational Research Asscciation (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Achievement Tests; *Anxiety; Educational Accountability; Elementary School Teachers; Evaluation Methods; Grade 3; *Information Utilization; Primary Education; *Standardized Tests; State Programs; Student Testing; *Teacher Attitudes; *Test Results
IDENTIFIERS Construct Validity

ABSTRACT

The meaning of standardized achievement testing to 50 third-grade teachers was investigated through a construct validation procedure using questionnaires and structured open-ended interviews. Subjects were from 18 schools performing distinctly above or below prediction on the Delaware Educational Assessment Program Longitudinal Study. Multimethod-multitrait procedures show the construct, evaluation acceptance, as valid, consistent, and distinct from trait anxiety. The meaning underlying the construct has dual evaluative dimensions, unanimously salient potency dimensions, and a deficient knowledge base. Implications are proposed for teacher education, state assessment models, administrative adjustments, and research on causal influences of teacher attitudes on testing. (Author/CTM)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Rita M. Fillos

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED161942

# EVALUATION ACCEPTANCE IN ELEMENTARY SCHOOL TEACHERS:

## A Construct Validation and Description of the Meaning of Standardized Achievement Testing

Rita M. Fillos
A. Jon Magoon

*University of Delaware*
*Newark, Delaware*

Prepared for presentation at the Annual Meeting of American Educational Research Association -- Toronto, Canada; 1978.

*Printed in U.S.A.*

TM007 977

Session # 25.13

# EVALUATION ACCEPTANCE IN ELEMENTARY SCHOOL TEACHERS

Although strong arguments against standardized achievement testing have been presented in the educational literature over the past few years, the influence of these tests is not diminishing. In fact, the use of standardized testing is basic to several growing state assessment programs (e.g., Pennsylvania, New Jersey, Maryland, Delaware and Michigan).

The impact of the primary role state assessment programs are assuming in operationalizing educational accountability, is drawing additional attention to the results of these assessments. Traditionally, formal assessment data has been little used by school decision makers (House, 1973b). However, the State of Delaware, for example, by making training in test administration and interpretation a requisite for certification as a school principal and by providing district grants for the study and implementation of assessment testing, has clearly demonstrated its intention of having assessment data used.

Purpose:

The primary purpose of this study is to measure the attitudes of selected Delaware elementary school teachers toward standardized educational evaluation, specifically toward the products of state assessment testing. In particular, the study attempts to validate a hypothesized construct, evaluation acceptance. A secondary purpose of the study is to assess the explanatory power of organizational and socialization conditions, as perceived by teachers, in accounting for variation in teachers' evaluation acceptance.

*Background:*

The evaluation of student performance by standardized achievement tests reflects and potentially affects teacher performance. Teachers are found to be highly sensitive to student evaluation procedures (Coughlin & Cook, 1974). Evaluation which is perceived as accurate and meaningful can be expected to produce job satisfaction if it indicates students are doing well and to cause great uneasiness if it indicates students are doing poorly. Therefore, successful implementation of standardized achievement testing programs is highly dependent upon the teachers' perceived meaning of the processes. When measures are used for a purpose that has little basis in the teachers' experience, they have no way to determine their meaning or assess their accuracy, and hence; the teachers comprehend that they have not been properly considered. Teachers, like other people in a threatening climate, are likely to become defensive and entrench. This is illustrated by the early reactions to proposals for accountability. Previous studies primarily considered the problems of testing insofar as they affect the students. Evaluation programs, however, either implicitly or explicitly reflect on teachers and other responsible school personnel.

Although its presence is clearly present in the protest literature against educational accountability (Cooper, 1972; Pratte, 1972; House, 1973; Landers, 1973; Nash & Agne, 1973; Bundy, 1974) the sociopsychological meaning of achievement testing in relation to teachers has received minimal empirical investigation. Research on the meaning of a concept, as argued by Messick (1975), should enter on construct validation. "Construct validity is the process of marshalling evidence in the form of theoretically relevant empirical relations to support the inference" (p.955). Research on any consistencies in the role

of a test for communicating the learning situations to the teacher in terms of her effectiveness is scarce. Some information on teacher attitudes, dissonance, conformity, and social comparison problems related to testing are provided in the ethnographic study done by Smith and Geoffrey (1968), the sociological study by McPherson (1972), and a few teacher surveys (Wolf, 1971; Wagner & O'Hanlon, 1968; Cicourel, et al., 1974).

Also, an intensive empirical study of school practitioners by Lortie (1975) deals with "the patterns of orientations and sentiments which are peculiar to teachers". This study provides a view of teachers who, as a result of their occupational socialization processes, are characterized as an occupational group strongly directed by conservatism, individualism and presentism (an emphasis on the here-and-now). Evidence for the presence of an orientation is established by the accumulation of independent instances of occurrence rather than by, say scores on a test for "conservatism". Independent instances of occurrence of "conservatism" would be demonstrated by a teacher's comment that everything that she knows that is worthwhile came to her by experience and, in another context, by the statement that "children don't care about learning as much as they used to." The orientations appear as an accumulation of informally expressed attitudes.

These orientations have been used to explain teachers' sources of rewards, teaching philosophies, career uncertainties, and the logic of their thoughts and feelings. Whether these basic orientations can be used to explain the responses of school personnel to a specific problem like evaluation needs further exploration. This model of aggregated self-report and interview evidence

appears to offer an opportunity to supplement a numerical measurement of the construct with documentation of its underlying vestiges in everyday life.

*Rationale:*

In this study the meaning of standardized achievement testing for teachers is investigated through two approaches: (1) the traditional multimethod-multitrait construct validation procedure; and (2) a less common open-ended interview procedure. Although dependent upon each other in methodological respects, e.g., the interview provides a second "method" necessary for the multimethod-multitrait analysis and the multimethod-multitrait establishes the technical relationships of the constructs focused on in the interviews, the two procedures are conceptually orthogonal or independent.

The multimethod-multitrait procedure operationalizes the teachers' behavior in terms of response scores on reliable questionnaire and interview scales to demonstrate a prespecified consistency in behavior, specifically, that if the construct exists it will be measurable in more than one form. However, the persistance of the failure of measurement technology to demonstrate difference in, say, the best ways of teaching, suggests that the notion of education as a unilateral, unidirectional process may be insufficient. The interview procedure, therefore, is intended to supply information required by a broader research philosophy which assumes that internal mental processes are major underlying determinants of behavior and that the underlying assumptions are important to research as the behavior. Therefore, the plan for this study is first to show that a construct of Evaluation Acceptance (EA) exists within elementary school teachers, and second, to describe some of the meaning and experiences that support its existence. The concern is with what the

construct represents, the understandings and perceptions that are reflected because teachers' characteristic beliefs have pervasive effects on their behavior, influencing the learning environment and consequently the children (cf. Halperin, 1976).

As mentioned previously, evaluation in the guise of accountability and standardized achievement tests has been consistently resisted by teachers. The general philosophical explanations for this behavior -- humanistic educational principles or a basic human fear of judgment, offer no heuristic value to a situation that needs clarifying. However, recent interview and ethnographic studies have shown relatively unique occupational and socialization characteristics which suggest observable situations that may be shown to be saturated with evaluation and which are empirically testable for consistency. For example, the reward system of elementary teachers has been found to be primarily psychic, individualistic, and present-oriented (Lortie, 1975; Sarason, 1971; Jackson, 1968); therefore, they can be expected to reject an evaluation system based on what they see as unproven technology, applied through secondary parties, temporally removed from the teachers' influence. Further, through the "slippage" in the loosely bureaucratic structure of schools (Bidwell, 1965) teachers have been able to hold to a source of power in their position as the primary evaluator of student performance (Myers, 1973) and for this reason also it is predictable they would wish to maintain as much autonomy as possible and would reject evaluation through achievement testing.

In addition to reward and power relationships with testing, the lack of a clearly identified technology of teaching, the vagueness of teacher evaluation procedures and a tradition of closed-mouth coping can produce uncertain-

ty among teachers and, with the lack of a reassurance structure, uncertainty can produce anxiety. Thus, the personality construct of anxiety, and the combined conservatism, individualism and presentism of these teachers are examined for their relationships with evaluation acceptance. The research literature displays the existence of conservatism, individualism and presentism and these orientations point to the probable presence of both anxiety and evaluation resistance. In order to clarify whether these are separate constructs, measures of both are included in both forms of the study to provide a source of divergent validity for the evaluation acceptance/resistance.

*Methodology:*

As a consequence of the preceding background and rationale, the following questions are addressed:

1. Can Evaluation Acceptance be measured and validated as a construct?
2. Is Evaluation Acceptance related to the teachers' traits of conservatism, individualism and presentism?
3. Is Evaluation Acceptance related to how "anxious" the teacher is?
4. Is Evaluation Acceptance related to how well a school performs as measured by the state evaluation test?
5. Do open-ended interviews provide a worthwhile methodology for developing the meaning behind a construct?

*Subjects:*

The subjects of this study consist of three male and 47 female teachers from eighteen elementary schools representing thirteen different school districts in Delaware. These schools were selected on the basis of a residual of at least 1.3 standard errors above or below predicted performance in mathematics and reading on the 1972-1975 Delaware Educational Assessment Program

Longitudinal Study. Twenty-two teachers are from schools which were below their predicted mean. The classification of school performance was not known to the researcher until all data had been collected. The questionnaires were administered personally to these 50 teachers to insure 100% participation and to build an impression of the importance of the study in order to insure the teachers voluntary participation in the time-consuming interview portion of the data collection.

A subsample of twenty teachers was selected for the interview portion of the study. Teachers were selected to be interviewed if their score on the EA scale of the questionnaire was one of the ten highest or ten lowest.

*Instruments:*

Four instruments were used in the study. All subjects received Form X2 of Spielberger State-Trait Anxiety Inventory (STAI) and the Evaluation in Schools Survey (ESS) prepared for this study. In addition, the teachers in the personal interviews were administered the Interpersonal Dimension of the S-K Inventory of General Trait Anxiousness (Endler & Okada, 1975) plus a list of thirty-five interview questions prepared to ascertain the concept structure or "meaning" of the data collected by the ESS. (Copies of interview questions and ESS forms are attached).

The ESS contains thirty input items comprising four categories: (1) demographic information, (2) perception of school evaluation policies, (3) Evaluation Acceptance and (4) conservatism, individualism and presentism (CIP).

In the course of the literature reviews, special note was taken of circumstances which appeared to invoke EA and/or CIP. A brief personalized restatement of this condition became the item to which the subjects were asked to agree or disagree. Other items grew directly from SA testing concerns.

This Likert-type instrument was pretested, refined by factor analysis and piloted on spring and summer graduate students (N=220) with the final version attaining a Pearson product moment correlation of .94 as test-retest reliability for the EA scale and r=0.90 for the CIP scale.

The interview script was designed to probe the teachers' value and concept structures surrounding CIP and EA and to reveal the teachers educational socialization experiences. The interview questions were carefully developed according to criteria listed by Bussis, et al (1976), pre-tested and sequenced. Although specific questions were asked, the purpose of the interview was not primarily to obtain "answers" but rather to allow subjects to speak freely about what the topics meant to them.

Reliability estimates of the interview questions consisted of the interrater reliabilities computer from three judges' estimations of the presence of EA and CIP traits in the tape recorded interview responses. The average interrater reliabilities for the EA and CIP scales are, respectively: 0.94 and 0.84

Validity estimates for these scales were obtained by the Product Moment correlation of the subjects' interview EA and CIP scores with her questionnaire EA and CIP scores respectively. These validity estimates are: EA, r=0.65 and CIP, r=0.27.

*Analytical Procedures*:

Three procedures were employed to analyze this data. The first is the widely used technique developed by Campbel & Fiske (1959) for evaluating the convergent and discriminent validity of concepts. This approach represents a multimethod-multitrait matrix analysis based on the rationale that each test employed is a trait-method unit; that is, that responses on the measure contain variance contributors due to the method of measurement as well as aspects of the construct being measured. Convergent validity is demonstrated

by the correlation of independent measures of the construct -- in this case those taken by interview and by questionnaire. For demonstration of discriminant validity it was necessary to use the same methods to measure at least one other construct. As it is suggested that the conditions of organizational ambiguity and socialization pressures which presumably lead to EA might also be expected to support trait anxiety, the problem is to determine whether EA is distinct from anxiety. The MM-MT process is strengthened if more than one alternative trait is included. Since the presence of EA is inferred from descriptive evidence of CIP, a question arises whether EA may be merely another manifestation of this orientation. Therefore, a measure of CIP is included in both the questionnaire and interview methods of data collection.

In addition to the convergent-discriminant validity approach, the establishment of a nomological network relating the characteristics of a proposed construct to each other or to other constructs strengthens the validation of a construct. Construct validation is enhanced when relationships between it and other variables can be predicted. For example, Coughlan and Cook (1974) find evidence to suggest that teachers in relatively high or low performance schools may contain teachers who are tightly directed, defensive and conservative. This condition suggests a testable hypothesis that a significant portion of EA variability may be contributed by the school's performance classification. Other suggested hypotheses involve the relationships of evaluation acceptance and perceived administrative support, source of rewards and length of teaching experience.

A third analytical procedure is used to supplement the deficiencies of a strict structural functionalism framework. This paradigm is concerned with

the quality of experience and the meaning as well as the occurrence of behavior. Thus, an indepth interview was employed as a strategy to solicit internal assumptions, experiences and beliefs about standardized achievement testing. The interview data were utilized in two ways. The first use provided raw material for judgments of EA & CIP scores necessary for MM-MT analysis. Additionally, the transcripts of the interviews were content analyzed to build aggregated narrative descriptions which describe the meaning that SA testing has to elementary teachers. The content data were also examined for consistencies in unsolicited comments and for anomolies that needed explanations.

*Findings and Discussion:*

Through multimethod-multitrait matrix analysis of the intercorrelation of the scores of three postulated constructs by two separate methods, Evaluation Acceptance is shown to have strong convergent validity. (See Table). It is also distinctly divergent from Anxiety and mildly conversely related to CIP, however, the CIP construct does not appear to be unequivocably stable across measurement methods. Evaluation Acceptance is therefore found to be a valid, consistent, measurable personality trait in elementary school teachers and this trait can be considered to be distinct from trait anxiety.

Regression analysis with dummy coded categories supplied the following understanding of the relationship of Evaluation Acceptance in a network of theoretical relevant variables.

School Performance classification is significantly related to the teacher EA scores with a coefficient of determination of 0.1033, $F=5.53$, df 1,48 significant at $p < .025$. The interview data revealed that teachers from high performance schools on the DEAP classification are more than twice as likely to accept SA testing as teachers from low performance schools.

From the literature on teacher characteristics, there appears to be justification in predicting that teachers who are more satisfied with general evaluation procedures in their school will also be more accepting of SA testing (as another form of evaluation). However, the findings of this study did not substantiate that relationship (r=.102). This may be due to the measures used because from the interview scripts it appears that the lack of relationship should probably be attributed to the SA testing base that this study used as a referent for evaluation. For most teachers, SA testing did not appear to be a salient organ in the perceived evaluation program. This evidence occurs in three forms: (1) teachers believe they are not evaluated by SA testing, but by the principal, (2) students are ostensibly not evaluated by SA testing, but by the teachers' judgment, and (3) teachers do not have a clear understanding of any evaluation process within their schools. Given that teachers say they believe that they are not currently evaluated on the basis of SA tests, it is a reasonable finding that EA, which is based on SA testing, is independent of the teacher's satisfaction with present evaluation methods.

Experience is reported to be the major means by which teachers believe they learn their job. Based on this socialization phenomenon, it was postulated the Evaluation Acceptance if it exists as an occupationally construct or trait should be related to a teacher's years of experience. From the data of this study, the number of years a teacher has spent in one grade seems to be more systematically important in certain attitude formation than total years as a teacher. After a natural log transformation, on the data from the 50 teachers sample to diminish skew in the distribution of years in third grade no linear relationship is found between experience and EA but there is a significant curvilinear relationship (eta=.23). There appears to be a ceiling

effect on the benefit of experience on attitudes at around five years in third grade; after that time attitudes toward EA tend to regress toward the mean. While attainment of advanced degree is not related to years taught in third grade, it does have a weak correlation ($p < .10$) with EA. With respect to experience, the findings show that for the first five or six years in a grade level, as experience increases so does EA, and that for teachers with any level of experience their acceptance of EA testing will tend to be higher if the teacher has advanced education.

Given that the ultimate goal of education is performance, indicated in this study by measures of student achievement, four variables in this study are found to relate directly to whether a school does better or worse than predicted by the State Assessment Test (DEAP) and a fifth is indirectly related through its association with ten of the direct variables. Multiple regression procedures are applied to build an explanatory model of the relationship of these variables for improving the accuracy of the State Models Prediction of School Performance (SCPR).

Wainer (1976) and Dawes and Corrigan (1974) have advanced a system of prediction using unit weighting of the variables in the linear model. This technique is shown to give results very close to those found by regression weights and have the advantage that they do not "use up" any degrees of freedom in their estimation. The present study benefits from that procedure by being able to use all five predictors without "overfitting" the model. The above authors show also that the unit weighting process is more robust for replication of the prediction using data from other samples. Typically, the $R^2$ in a validation sample will show shrinkage from the $R^2$ in the original sample due to the overfitting of the original model to fit the outliers in the data into the estimates of the regression weights. Using the unit weighting

system, the weights are the same for both samples so shrinkage of any appreciable size does not occur.

The predictors in this study had been selected for theoretical reasons and maintained in the equation because their correlations were statistically significant with the dependent variable. The process for combining them into one composite predictor for this analyses is as follows: The scores were standardized and then added up. The correlations of EA, EVLEVR, YRSTRD and ADVDGR with SCPR were positive so these scores were summed. The z score for STAI was subtracted because its correlation with SCPR was negative. The formula for the new predictor variable is

$$\text{Unit Weight Predictor}=Z_{ea}+Z_{evlevr}+Z_{yrstrd}+Z_{advdgr}-Z_{stai}.$$

The results of this regression, taken in two steps are:

| Dependent Variable | Independent Variable | $R^2$ | F |
|---|---|---|---|
| Standard Error of School Performance | $Z_{evl}+Z_{yrstrd}+Z_{adg}-Z_{stai}$ | 0.25 | 15.96*** |
| | $Z_{ea}$ | 0.30 | 9.86** |

F test of increment in explained variance due to addition of EA

$$F_{1,47}=16.66***$$

Therefore, by utilizing a unit weighted multiple regression procedure, the error of prediction in the State Model can be reduced by 30% by including individual (not aggregated) measures of teaching experience at the current grade level, trait anxiety, advanced degrees, satisfaction with general evaluation procedures and Evaluation Acceptance. The increment in variance accounted for by Evaluation Acceptance, entered after all other variables, is significant at $p < .001$ suggesting that this construct carries unique information over the other variables measured.

*Deductive Description:*

The major portion of the open-ended interview was designed to assess what meaning several organizational and occupational conditions brought to teachers' evaluation acceptance. These conditions were teacher perception of organizational control, reward sources, peer support and training adequacy. Additional understanding of teacher's EA levels comes from probing their expectations and knowledgeability concerning SA testing.

Intensive sifting and recombination of the information provided by the interview tapes yields the following background information which broadens the meaning of evaluation through SA testing to third grade teachers.

1. These teachers are aware that SA test scores have the potential to diminish the within-class power they consider central to their ability to function in the classroom.

2. Although SA tests are commonly dismissed as irrelevant, inaccurate and inappropriate, these teachers still feel rewarded if their students do well in them and chagrined if they do not.

3. These teachers are deeply child oriented. They perceive the moral and affective domains as the most important portion of their teaching responsibility. Thus, SA tests dealing with only the cognitive domain are a gap in the teachers mediation of the child's growth. They do not come from teachers and do not always provide assistance to aid the teacher with individual children's needs.

4. The expectations that teachers hold for SA tests are at cross purposes with the traditional intentions of SA test batteries, which are constructed to provide broad, long-term evaluation of curricula, program and staff.

5. Teacher knowledge of SA test capabilities is informal and sketchy. Many say they see no value for testing to anyone but teachers and suspect a "hidden agenda" of testing is held by administrators.

6. Strong professional self-assurance in a teacher can flatten the intensity of problematic situations including those of evaluation through student testing.

*Conclusions and Implications:*

The primary purpose of this study was to describe the meaning of standardized achievement testing to elementary school teachers and to validate the construct of Evaluation Acceptance which represents this meaning. The tasks necessary to achieve this purpose were measuring and validating the construct, investigating its relation to other variables involved in student achievement, and describing the attitudes underlying the construct which were revealed through teacher interviews.

*Validity of Evaluation Acceptance:*

Application of the Campbell and Fiske procedure clarified that EA has the necessary convergent and divergent validity to qualify as a unique valid construct. The negative direction of the correlation of the anxiety measures is an unexpected finding, but the SR survey form is a rather personally revealing experience, and a socially anxious person might be expected to be "defensive" (Sarason, 1960) which would lead the subject to adjust the responses to research those considered most socially appropriate. The more anxious on STAI the more adjustment on SR, hence the negative correlation. Although there is some evidence that EA is negatively related to orientations of CIP any conclusions must be held until CIP is validated more strongly. While the reliability of the CIP measures within each method was reasonable, the scales did not correlate across methods. This very likely is a measurement problem of a unidimensional scale. The questionnaire, for example, may have been more sensitive to "individualism" whereas, the tape judges may have been more alert to "conservatism". The CIP scale consists of abstract orientations which do not have the advantage held by EA of being tied to a concrete experiential representation.

The importance of experience surfaced several times in this study. Over the first few years of teaching in third grade, as experience increased so did Evaluation Acceptance but a ceiling on this effect is reached around five years. By that time, the teacher apparently is totally socialized into that grade level routine, feels less threatened in the role, (may have tenure and be evaluated infrequently) and how he/she feels about SA tests depends on variables pertaining to the type of children and/or administrative situation such as having to manually correct and score the tests. However, it should be emphasized that total years teaching which certainly would reflect tenure was not significantly related to EA. Evaluation Acceptance canrot be predicted on the basis of total teaching experience, but by years at the grade level according to the data on this sample of fifty teachers.

The finding that variables proposed and measured in this study are able to account for nearly one-third of the deviation from prediction indicates their strong heuristic value for reexamination of the DEAP model. The accuracy of this multiple regression model is dependent upon whether or not all necessary predictions are included. Although the DEAP model presently includes measures of teaching experience and teacher advanced degrees, they are measured by averaging the average number of years of the teaching staff in the school for the years 1972-75. Use of individual rather than aggregated measures of these variables plus the teacher attitude measures on anxiety, general evaluation and Evaluation Acceptance can account for 30 percent of the deviation from prediction in the Delaware State Evaluation Assessment Program Longitudinal Study.

Considering the importance that the accuracy of this prediction has to school politics and administration decision-making, the model should be re-

fined wherever possible. Data on two of the tested new variable, advanced degrees and years at grade exist in school files and scores on the attitude measures could be obtained by administration of one questionnaire by the principal.

*Meaning of Evaluation Acceptance:*

The influence of organizational and socialization conditions on the meaning to teachers of evaluation through standardized achievement testing is clarified by employing Osgood, Suci and Tannenbaum's (1957) dimensions of meaning. If the major dimension of meaning is an evaluative one, the emphasis in a study of the meaning of EA would be on acceptance or rejection of SA testing activities. In this study, the mean score on the questionnaire EA scale suggested that these third grade teachers were definitely inclined to reject SA testing. But when this same scale was presented in open-ended interview form, an examination of the responses showed EA to be a construct with dual components rather than a bipolar continuum of like-dislike. Kerlinger (1967) noted that dualism of attitudes inhere in the criterial referents of the attitude. If a referent is criterial for an individual, it is relevant and significant for him. What is relevant or criterial to one group of individuals may not be to another. In the case of SA testing, the majority of the beliefs presented for discussion were recognized by all teachers. For some these were criterial for attitudes of evaluation rejection or acceptance, but for others the beliefs made no difference to acceptance or rejection at all. There appear to be some teachers for whom evaluation and SA testing in themselves simply are not criterial referents because they perceive the tests to have no effect on <u>their</u> behavior.

Which component of this dualism a teacher might fall into appears to be determined by lack of experience, low self-assurance, "problem children," and administrative pressure. Teachers who experience three or four of these conditions exhibited a range of EA scores. Teachers who found SA testing was not criterial to their own behavior were complacent about testing activities and there was much less variation in their EA scores.

The potency dimension is also salient due to concern for power relationships. The findings of this study demonstrate that more than a majority of the subjects believe that SA test scores have present or poential ability to diminish their authority as the classroom expert. In addition to the teachers' concern that administrators and parents may be more convinced by test scores than grades, the teachers found that they too were susceptible to feeling rewarded or punished by test results, even though they professed to find them neither valid nor helpful. The potency dimension, i.e., strength of status and control, was found to have a consistent rather than a dual or bipolar nature in this study, all subjects perceived this to be a salient area. This was probably a reflection of the political reality of the increased emphasis on SA testing and evaluation in Delaware and surrounding states.

Underlying the evaluative and potency dimensions of the meaning of SA testing to teachers is the requirement of a knowledge or information base. According to the findings from this study, this base was a misfit for SA testing in at least three aspects: goal concensus, test methodology and administrative support. For 85 percent of the interviewed teachers inculcating a moral obligation to attend to learning and abide by social rules was the core

of their goal behavior; only three teachers seemed primarily concerned with academic goals. Given this strong socio-emotional goal orientation, it is not surprising that the norms of teaching content became a secondary goal to the norms of teaching the child and SA testing was often rejected because it seemed to have nothing to do with teaching children to get along with each other or to love learning. Many teachers mentioned a time squeeze which increased the difficulty of "really getting to know the children". Thus, they seriously resented activities which were perceived to be non-productive in advancing this child-teacher relationship. SA testing was one of these activities. Furthermore, both socio-emotional and academic goal oriented teachers complained about SA tests' shortcomings as a diagnostic tool for a child's particular abilities. The prevailing nature of this complaint demonstrated the teachers' lack of knoweldge of the goals of standardized achievement survey batteries as well as their minimal technical understanding of SA batteries' capabilities and limitations.

In addition to the lack of diagnostic helpfulness, SA tests were criticized for not being available for pre-post gain measures. This indicates another technical problem because with most SA tests the standard error of measurement may be larger than the gain in correct items which allegedly represent one grade equivalent of growth. Thus, an apparent gain or loss over the six month period which teachers would like to assess could not be distinguished from the measurement error of the test. On the positive side, SA batteries can be used by teachers to diagnose group or class weaknesses, to indicate students' relative strengths in different subject matter fields and to show growth patterns over the years, although very few teachers recognize these applications.

The function of SA tests is to help in decision making, but in this study it is apparent that most schools had not identified to teachers whose decision making the SA tests are to assist. Since very few teachers reported knowing why SA tests were given, it is speculated that the purpose must not be for teachers' use but to motivate administrative and supervisory efforts by affording an overall measure of the effectiveness of the school organization and of the prevailing administrative and supervisory policies. (The test results not only have the potential of evaluating teacher performance via student achievement but this can be taken a step further to evaluate the administrations' performance via the teachers' performance via the student achievement). Since the tests did not fit the teachers' information needs and since the administration did not inform them as to what other expectations were held for test results, the tests are a source of discord subject to superstitions and suspicion brought in to fill the knowledge vacuum.

*Usefulness of Interview Methodology:*

Part of the purpose for including the interview technique in the study was to investigate its usefulness for illuminating the meaning of the proposed construct. Much of the problem introduced by the subjects' concern for providing sociably desirable responses can be relieved by low-keyed interviewing techniques, carefully worded questions and promises of total confidentiality. A distinct advantage is demonstrated because discussions of interview procedures center around participants while discussion of questionnaire information centers on items or variables. While knowledge of the relationship of the variables is necessary to understanding the phenomenon, it is the belief structure that is essential to understanding the interactive nature of the persons and the phenomena. While the interview information did not always clarify the

situation, it added texture to the construct and sometimes suggested otherwise unanticipated research directions.

*Implications for Teacher Education:*

If public education is going to compete successfully for a share of a communities' moral and financial support, accountability advocates insist that it must increase its ability to demonstrate fulfillment of perceived learning goals. Ideally, this competency demonstration would be convincing to persons inside the schools as well as those it services. In view of the finding in this study, that agitation rather than satisfaction is a large component of the meaning of SA evaluation to many teachers in Delaware, it seems that elimination of some or all of this malaise could be expected to improve teacher morale and lead to more effective use of their talents. Abolishing testing, of course, might remove these irritations but would be incompatible with the goals for demonstrated competency. SA testing programs, while flawed, are ingrained in the traditions of school evaluation and, furthermore, are the most acceptible evidence to outside policy makers as well as educators. SA emphasis shows little signs of waning. In view of this situation, the findings of this study imply that for pragmatic as well as theoretical reasons, teachers should receive the training necessary for them to live with SA testing more comfortably.

Inasmuch as teachers perceive a dubious efficacy when administrators pay for SA testing and don't show that it is utilized and, furthermore, the teachers are aware that these test scores have potential effects on their power base, school principals need to be convinced of the need for straight thinking and clear in-school communication about testing programs. At the same time, there are implications that principals should display to both teachers

and parents how teacher evaluation of students supplements formal survey testing and how all tests build into a student evaluation program.

Teachers report sharing ideas about materials or activities but evidence of a community of sharing with respect to testing experiences was not observed in this study. What evidence does exist seemed to indicate that teachers did not talk about tests among themselves because (1) teachers are powerless to effect SA testing and (2) there is a social stigma attached to conscious effort to reach for high SA test performance. This situation suggests a need for the administrator to be the distributor of useful test knowledge which originates within the school or collected from outside sources such as state, district or research reports. His positional authority could enhance the respectability of examining testing successes with as much attention as is given to failures.

*Implications for Methodological Improvements in Future Research:*

Although all interview items provided useful information, those directed toward understanding of organizational influence did not elicit particularly deep answers; thus, it is not clear whether the teachers were being wary about answering questions in this direction or whether they actually had not thought much about their position in an organizational network. Therefore, some more stimulating interview items relating to teacher relationships with administrators and peers need to be developed. Perhaps increased use of personal experience probes rather than "do you think" questions would be helpful here. It is also possible that conducting the interviews away from the school might also improve teacher openness and permit them to give undivided attention to the interview.

Observational and interview studies have often been criticized from the standpoint of replicability. Goodlad (1977) noted that it was surprising and disappointing, given the significant and provocative nature of some of the findings of these kinds of studies, that there was seldom further inquiry to follow up and extend the conjectures. Additionally, in the absence of a random selection from a large number of possible schools, replication seems to provide the best indication of generalizability of findings from interview and construct validation studies. The accumulation of similar evidence from divergent sources lends generalizability to the status of many findings across the studies. This particular study replicated interview responses from both Lortie (1975) and Jackson (1968) interviews with respect to occupational choice and teaching goals. It diverged from Lorties' findings on the isolation of teachers reported in the late Sixties, but agreed with the descriptions collected by Bussis, et al. (1976) in the mid-Seventies, viz., that more cooperative interaction was desirable although usually not practiced on a day-to-day basis. Therefore, future research of this topic might benefit by attempting to (1) replicate findings of this study from teachers of other grades and locales and (2) replicate portions of other interactive studies relevant to the background and meaning of evaluation to teachers in elementary schools.

*Implications for Future Research:*

The EA scale developed and validated in this study provides an entry for examination of school testing in at least two directions. Given that teachers can be categorized into those who are SA test accepting and others who are SA test resistant, the question immediately arises: What differences, if any, does this make in their classroom behavior? Observation studies of teachers preparing their students and actually administering the tests might provide an answer to this question; however, in addition to the usual prob-

lems of gaining access to classrooms, this type of study would be hampered by the possibly legitimate complaint from the schools that an observer would disturb the teachers. Therefore, this kind of information might be best collected as part of an overall ethnographic study of classrooms by an observer who was present also at non-testing times and who had been alerted to give careful attention to testing related incidents.

Another worthwhile study applying the EA scale could involve a two-stage experimental study to examine the causal relationship of EA to student performance on SA tests. The first stage would involve categorizing teachers by EA score, then over a period of several months retraining the teachers regarding their testing expectations and/or enhancing their self-confidence, and then reassessing their EA scores for evidence of change. The second step of the study would be to look at the pattern of achievement for the teachers' students in relation to whether the teacher's EA increased or decreased. The effects might be small over a short time; therefore, the design would need to be as powerful as possible, removing as many sources of confounding as could be measured from the error term: IQ, previous grades, school, teacher experience and other effects known to contribute systematic variance to achievement.

The teachers in this third grade sample provided very little information about the principal's contribution to their understanding of testing. Thus, the question of meaning of SA testing might be investigated by a study parallel to this one involving elementary principals rather than teachers. Implementation of strategies within the schools to change teacher EA would be impeded without an understanding of the meaning of SA evaluation to principals.

*Summary:*

The results of this study indicate that an elementary school teacher holds a consistent attitude toward evaluation and toward standardized achievement testing as the agent for this evaluation. This attitude can be measured and each teacher assigned a rating on an evaluation acceptance/resistance scale. Evaluation Acceptance (EA) is shown through three types of analysis to be a valid construct or personality trait in these third grade teachers.

Although there is a statistically significant correlation ($p < .05$) of the scores on this EA scale with the schools' performance on the State Assessment test, there is no strong evidence that knowledge of school standing directly effects EA. Teachers report that they do not recognize the State Assessment test as an accurate indicator of school performance and consequently of their own performance. However, when EA is used as a predictor in the State's Model the error of prediction on how a school will do in reading and in math is significantly ($p < .001$) and substantially reduced. This suggests that teacher attitudes can contribute to overall student performance on standardized achievement tests.

Standardized achievement testing as the experiential referent for the Evaluation Acceptance attitude does not produce a bipolar continuum. To some teachers it has such a small personal importance that they form no strong attitudes either way -- for or against SA testing. These teachers are the minority. Generally, teachers show a variety of evaluation acceptance levels. The attitude is related to their beliefs about testing purposes, their perceptions of their principal's support and the strength of their personal self-concept.

Organizational conditions such as clarification of the use of testing in the contemporary teacher's role appear to be as influential for evaluation acceptance as socialization conditions (e.g., increasing the teacher's perception of testing as a source of approval as well as a source of censure) and are probably more adjustable. Additionally, more training on testing capabilities and specific limitations would go far in strengthening a teacher's acceptance of standardized testing and consequently, of evaluation through standardized testing.

## REFERENCES

Bidwell, C. E., The school as a formal organization. In J. G. March (ed.), Handbook of Organizations. Chicago: Rand McNally, 1965, pp. 972-1022.

Bundy, R. F. Accountability: a new Disneyland fantasy. Phi Delta Kappan, 56(3), November 1974, 176-180.

Bussis, A. M., Chittenden, E. A., & Amarel, M. Beyond Surface Curriculum: An Interview Study of Teachers' Understandings. Boulder, CO: Westview Press, 1976.

Campbell, D. T. & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Cicourel, A. V., Jennings, K. H., Sybillyn, H. M., Lecter, K. C., Mackay, R., Mehan, H., & Roth, D. R. Language Use and School Performance. New York: Academic Press, 1974.

Cooper, L. G. Decision ability, not accountability. Journal of Higher Education, 1972, 43, 655-660.

Coughlin, R. J., & Cook, R. A. Work Attitudes. In H. J. Walbert (ed.), Evaluating Educational Performance: A Sourcebook of Methods, Instruments and Examples. Berkeley, CA: McCutchan, 1974.

Dawes, R. M. & Corrigan, B. Linear Models in Decision Making. Psychological Bulletin, 1974, 81(2), 95-106.

Endler, N. S. & Okada, M. A multidimensional measure of trait anxiety: the S-R inventory of general trait anxiousness. Journal of Counseling and Clinical Psychology, 1975, 43(3), 319-329.

Flanagan, J. C., et al. The American High School Student. Pittsburgh, University of Pittsburgh, 1964.

Foskett, J. M. The Normative world of the elementary school teacher. Eugene Oregon: The Center for the Advanced Study of Educational Administration, 1967.

Halperin, M. S. First-grade teachers' goals and children's developing perceptions of school. Journal of Educational Psychology, 1976, 68(5), 636-648.

House, E. R. The price of productivity: who pays? Today's Education, 1973 62(6), 65-69.

House, E. R. School Evaluation. Berkeley, CA: McCutchan, 1973.

Jackson, P. W. Life in Classrooms. New York: Holt, Rinehart & Winston, 1968.

Kerlinger, F. N. Social attitudes and their criterial referents: a structural theory. Psychological Review, 1967, 74, 110-122.

Landers, J. Accountability and progress by nomenclature: old ideas in new bottles. Phi Delta Kappan, 1973, 54(8), 539-541.

Lortie, D. C. Schoolteacher: a Sociological Study. Chicago: The University of Chicago Press, 1975.

McPherson, G. H. Small Town Teacher. Cambridge, MA: Harvard University Press, 1972.

Messick, S. The standard problem: meanings and values in measurement and evaluation. American Psychologist, 1975 (October), 955-966.

Myers, D. A. Teacher Power: Professionalization and Collective Bargaining. Lexington, MA; Lexington Books, Heath, 1973.

Nash, R. & Agne, R. The ethos of accountability. Teachers College Record, 1972, 73, 357-370.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. Measurement of Meaning. Urbana, IL: University of Illinois Press, 1957.

Pratte, R. An uneasy inquiry into accountability. Intellect, 1972, 101(2343), 37-40.

Sarason, S. B. The Culture of the School and the Problem of Change. Boston: Allyn & Bacon, 1971.

Sarason, S. B. Lighthall, F. F. Davidson, K. S., Waite, R. R., Ruebush, B. K. Anxiety in Elementary School Children. New York: John Wiley, 1960.

Smith, L. M. & Geoffrey, W. The Complexities of an Urban Classroom: An Analysis toward a General Theory of Teaching. New York: Holt, Rinehart & Winston, 1968.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologist Press, 1970.

Wagner, R. & O'Hanlon, J. Teacher attitude toward evaluation. Journal of Teacher Education, 1968, 19, 471-475.

Wainer, H. Estimating coefficients in linear models: it don't make no nevermind. Psychological Bulletin, 1976, 83(2), 213-217.

Wolf, R. L. The role of the teacher in classroom evaluation. Unpublished Master's thesis. University of Illinois, 1971.

MULTIMETHOD-MULTITRAIT CORRELATION MATRIX

| Questionnaire | | | | Interview | | |
|---|---|---|---|---|---|---|
| | EA | CIP | STAI(Anxiety) | EA | CIP | S-R(Anxiety) |
| EA | (.93) | -.14 | -.10 | -.65 | -.34 | -.16 |
| CIP | -.14 | (.80) | .01 | .49 | .27 | .18 |
| STAI | -.10 | .01 | (.77) | .10 | -.24 | -.50 |
| EA | -.65 | .49 | .10 | (.94) | -.09 | -.17 |
| CIP | -.34 | .27 | .24 | -.09 | (.84) | .51 |
| S-R | -.16 | .18 | (-.50) | -.17 | .51 | (.77) |

Questionnaire

Interview

TABLE I

## SURVEY ON EVALUATION IN SCHOOLS

The following questionnaire has been developed to gain some needed insight into the perplexing problem of the meanings attached to school evaluation. Please read the statements, then place an X between the vertical marks on the line at the point which describes how you feel about the message in the statement. PLEASE RESPOND TO ALL ITEMS

A. The first group of statements covers a variety of topics related to achievement tests and educational evaluation.

1. Student gains on achievement test scores give an excellent measure of teaching effectiveness.

| Strongly Disagree | Disagree Somewhat | Agree Somewhat | Strongly Agree |
|---|---|---|---|

I I I I I I I I I I I I I I I I I

2. After I have taught a child for several months, I can judge his ability better than any test can measure it.

I I I I I I I I I I I I I I I I I

3. Achievement test results can give more useful information to me as a teacher than they can give to administrators.

I I I I I I I I I I I I I I I I I

4. I have a clear picture in my mind by which I judge "success" in each learner.

I I I I I I I I I I I I I I I I I

5. In measuring the improvement in performance gains of my class, it seems unfair to include students who will never learn or who are rebels.

I I I I I I I I I I I I I I I I I

6. A good thing about achievement tests is that even when the gains they show are small, they are concrete.

I I I I I I I I I I I I I I I I I

7. Standardized achievement tests need to recognize that it takes some children longer to learn a topic than others.

| Strongly Disagree | Disagree Somewhat | Agree Somewhat | Strongly Agree |
|---|---|---|---|

I I I I I I I I I I I I I I I I I



©1974

8. I find that standardized achievement tests reflect fairly accurately what the students really know.

| Strongly Disagree | Disagree Somewhat | Agree Somewhat | Strongly Agree |
|---|---|---|---|

I I I I I I I I I I I I I I I I I

9. I dont feel there is a need to report the results of one students performance in terms of how other students have achieved.

I I I I I I I I I I I I I I I I I

10. Since teachers work under different conditions (i.e., the students they receive, the curriculum objectives, etc.,) I feel that teachers should not be judged by any set standard.

I I I I I I I I I I I I I I I I I

11. The methods I use are based more on what I have learned through experience than on what I learned in college courses.

I I I I I I I I I I I I I I I I I

12. Because school goals are vague, I generally interpret them myself in order to implement them into my teaching.

I I I I I I I I I I I I I I I I I

13. The person who can give the most knowledgeable evaluation of my job performance is myself.

I I I I I I I I I I I I I I I I I

14. The structured guidance I give my class makes a difference in my students' achievement gains.

I I I I I I I I I I I I I I I I I

B. The following group of statements is related directly to standardized achievement testing.

15. Standardized achievement testing is a major influence on my day-to-day teaching activities.

I I I I I I I I I I I I I I I I I

16. Standardized achievement tests <u>can</u> be used to show how well students have learned.

| Strongly Disagree | Disagree Somewhat | Agree Somewhat | Strongly Agree |
|---|---|---|---|

I I I I I I I I I I I I I I I I I

17. Standardized achievement tests <u>should</u> be used to show how well students have learned.

| Strongly Disagree | Disagree Somewhat | Agree Somewhat | Strongly Agree |
|---|---|---|---|

I I I I I I I I I I I I I I I I I I I

18. Average test scores for the class reflect teacher effectiveness more than the individual's test scores.

I I I I I I I I I I I I I I I I I I I

19. In broad-scope tests such as standardized achievement tests differences in student backgrounds must be considered in evaluating their relative performance gains.

I I I I I I I I I I I I I I I I I I I

20. The standardized achievement tests <u>can</u> be used to show teaching effectiveness.

I I I I I I I I I I I I I I I I I I I

21. Standardized achievement tests <u>should</u> be used to show teaching effectiveness.

I I I I I I I I I I I I I I I I I I I

22. I believe standardized achievement testing is important.

I I I I I I I I I I I I I I I I I I I

23. I believe standardized achievement testing is important but I'm not sure if school administrations can use the results non-politically.

I I I I I I I I I I I I I I I I I I I

24. I find that one of the directly useful things about standardized achievement tests is that, in a time when tough school decisions have to be made, it provides information which is equivalent for all districts.

I I I I I I I I I I I I I I I I I I I

25. If we had open discussion of the standardized achievement test results we might change the public's opinion of school personnel.

I I I I I I I I I I I I I I I I I I I

26. I believe that increased attention to standardized achievement scores will tend to create a narrow curriculum aimed toward the test items.

I I I I I I I I I I I I I I I I I I I

| Strongly Disagree | Disagree Somewhat | Agree Somewhat | Strongly Agree |
|---|---|---|---|

27. I don't think it benefits a teacher very much to give a lot of attention to standardized achievement tests because the results are generally not used for student improvement.

| Strongly Disagree | Disagree Somewhat | Agree Somewhat | Strongly Agree |
|---|---|---|---|

I I I I I I I I I I I I I I I I I

C. The final group of statements concerns your feelings about the school where you teach.

28. I think my work performance is judged fairly here.

I I I I I I I I I I I I I I I I I

29. Little effort is made here to evaluate the effectiveness of our instructional program.

I I I I I I I I I I I I I I I I I

30. The procedures for judging my performance are helpful to me in improving my work.

I I I I I I I I I I I I I I I I I

31. I'm rarely told whether I'm doing good work.

I I I I I I I I I I I I I I I I I

32. We lack satisfactory procedures here for evaluating student programs.

I I I I I I I I I I I I I I I I I

33. Even when you take into account differences in student ability other schools in this locality seem to be ahead of this one in educational effectiveness.

I I I I I I I I I I I I I I I I I

34. I would rate this district as one of the best, overall, for those who want to work in education.

I I I I I I I I I I I I I I I I I

D. Please furnish the following background information:

35. Number of years teaching this grade._____
36. Number of years teaching. _____
37. Level of academic preparation (B.S.,M.Ed., etc.)_____
38. Sex _____
39. Name_____

Thank you very much!

EVALUATION IN SCHOOLS INTERVIEW

1. On the Evaluation Questionnaire you recorded that you have taught about ___ years and that ___ of those years were at third grade level. How do you happen to be teaching third grade?
2. What other grades have you taught?
3. Did you always teach in this district?
4. Where have you received most of your training?
5. What were the major attractions that teaching held for you at the time you decided to enter that field?
6. In what ways has teaching been different from what you expected when you made the decision to become a teacher?
7. Of the various things you do as a teacher, what do you consider to be the most important?
8. To what extent are you free to do more or less as you think best? i.e., in what ways are you "your own boss"?
9. Do you think that the control of the way you work is more from within yourself (your goals, values, personality, etc.) or from factors outside like your work environment, school organization, etc.? e.g., Do you think you would be pretty much the same kind of teacher in any school or with any group of children or would you be likely to shift in each situation?
10. Could you take a position on the old argument of whether teaching is a science or an art and tell me what that position would be?
11. From the following six classes of values, which two probably are the most important you you: theoretical, social, political, religious, aesthetic, economic.
    Theoretical: concerns ideas, speculation, logic, reationality
    Social: concerns interactions with people
    Political: concerns influencing or guiding policy or control
    Religious: relates to religious beliefs
    Aesthetic: appreciative of or responsive to beauty
    Economic: relates to the production, distribution and consumption of goods and services.
12. From descriptions made in Sixties, teaching comes across as a lonely profession. Do you feel that is true now? Do you think an elementary school teacher is generally isolated from the presence and support of other adults? (at work, that is)

25. As far as you know, do the teachers and principal here seem to agree about most facets of standardized achievement testing?

26. Do you have any feeling that standardized achievement tests might be used to side-step the teacher as a decision maker? i.e., Is the assessment of student achievement from achievement tests likely to supercede the assessment made by classroom teachers?

27. If you were extremely unhappy about some aspect of standardized achievement testing in your classroom, what could you do about it or what would you want to do about it?

28. Can you remember why you said (agree or disagree) to "Student gains on achievement test scores give an excellent measure of teaching effectiveness?"

29. What were you thinking when you responded (agree or disagree) to "I find that standardized achievement tests reflect fairly accurately what the students really know"?

30. From the behaviorists' point of view, people are inclined to participate in activities that reward them. Can you see standardized achievement testing as a source of rewards to a teacher?

31. How do you feel about the suggestion that, since standardized achievement tests are not perfect, schools should seriously consider discontinuing their use until the tests can meet these criticisms?

32. Thirty years ago some schools used a standardized achievement battery but did not generally give the results much attention. How do you think education would gain or lose if we maintained that condition?

33. If you were to vote for continuation of standardized achievement testing in your school, what is the probability that you would vote "yes"?

34. What do you think that the evaluation of your work performance at this school is based on? i.e., Do you think that the core of your evaluation is genuinely based on your performance with the students and/or how much is based on the students' performance?

35. Are there any other points about standardized achievement testing that you would like to comment on? Are there some important questions that I have not asked?

13. Do you have opportunities to discuss your classroom successes and failures with your peers? Are you able to benefit from other teachers' experiences?

14. Would you consider mamy of the teachers in this school your "personal" friends?

15. In the studies mentioned earlier, the teachers described the sources of success they felt to be incidents with individual students. Does this agree with your sources of satisfaction from teaching? Do you think these incidents (if they occur) have much influence on your overall evaluation of students' progress?

16. Where do you feel standardized achievement testing fits into your job? i.e., Do these tests determine your goals, adjust your methods, evaluate your proficiency or anything else?

17. Do you perceive evaluation through standardized achievement testing as a threat or a challenge? Please expand on why you feel that way.

18. Do you feel that teachers should have a greater participation in the overall student evaluation process? If yes, how might this be obtained? e.g., Should teachers select tests, or be responsible for making decisions based on the test results, etc.?

19. What effect do you think increased usage of standardized achievement testing would have on the autonomy of the elementary school teacher? Do you think increased use of standardized tests would limit your independence in teaching in any way?

20. Have your attitudes toward standardized achievement testing changed in the last three or four years? Why do you think this is so?

21. Please try to describe what you expect to be the use of giving standardized achievement tests to your students. What do you expect to "get" out of standardized achievement testing?

22. How do you feel the accuracy of standardized achievement tests could be improved?

23. In what ways might the usefulness to teachers of standardized achievement tests be increased?

24. Can you think of some ways in which teacher training--college courses of inservice sessions--have or might have made you feel better prepared to work with standardized achievement tests?